EXHIBIT

F

ing the relative (and perhaps absolute) income of the previously highly-paid workers whose skills have been automated.

It is too early to determine with much certainty how this will play out for AI, whether the impact on any particular job will be positive or negative. Research is beginning to emphasize which jobs are most likely to be affected rather than lost (*8*). For example, that classification tasks such as image recognition can be done with AI will affect workers whose jobs involve classification tasks, such as radiologists (*7*). Recent work examining differences between generative AI (specifically, LLMs) and nongenerative AI [as described in (*7*)] shows that millions of jobs have the potential to be affected by LLMs. Notably, these studies emphasize that "affect" does not mean "replace." For many jobs, automating some aspects of the workflow might increase productivity, the wages of workers who have that job, and the number of workers hired to do that job.

Even when some jobs get automated, that might complement the tasks done by other workers. Many empirical exercises [for example, (*7*–*9*)] emphasize the direct impact on jobs, but they do not explore the jobs that might be enhanced through complementary production processes. For example, in January 2023, there were 186,417 job postings in the United States that specified language skills (such as Spanish Language or American Sign Language), or about 5% of the total job postings (see SM). Automating language translation would directly affect many of these jobs. At the same time, many other jobs that do not require language skills would also be affected. For example, a recent study showed that small businesses that used a rudimentary automated language translation tool on eBay experienced a 17.5% increase in exports to markets where that language is used (*15*). Automation of some jobs could create opportunities for those whose work would appear to be unaffected, as measured with the tasks and skills involved in current workflows.

CONCLUSION

Many economists who have studied the impact of automation on labor markets have argued recently that the direction of AI research needs to be changed away from automating tasks to focusing on overall job augmentation. The implicit argument is that a focus on augmentation will lead to more complementarity with lower-wage labor and more new tasks. However, many recent advances in AI that have been developed with the explicit goal of task automation have appeared to increase worker productivity; that is, task automation has been labor augmenting. Furthermore, AI technology may disproportionately augment lower-skilled labor, reducing income inequality. This, at the very least, calls into question whether a change in the innovator's mindset is needed: Task automation may be a path to substantially improved labor productivity.

This potential to reverse the recent trend toward skill-biased technical change does not mean that AI is without risk. Other concerns remain, including those related to privacy, liberty, democracy, and monopoly power (*3*). Our emphasis is on understanding that one person's automation is another's augmentation, and that it is difficult for engineers or policy-makers to pick which particular innovation will increase or reduce inequality overall. We believe that both regulators and engineers should be careful in shutting down a particular technology trajectory because it appears to automate human work. In the process of automating some work, other work can be augmented.

Often, our analysis suggests that such augmentation from AI will increase the job productivity of less-skilled workers who can now perform at levels achieved by their skilled counterparts. This suggests that skill premia that have contributed to widening inequality may be eroded. Thus, it is quite plausible that the use of AI to automate tasks will both increase productivity and decrease income inequality. If so, then we may want more automation, not less. ■

REFERENCES AND NOTES

1. E. Brynjolfsson, *Daedalus* **151**, 272 (2022).
2. J. S. Gans, A. Leigh, *Innovation + Equality: Creating a Future that is more Star Trek than Terminator* (MIT Press, 2019).
3. D. Acemoglu, S. Johnson, *Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity* (PublicAffairs, 2023).
4. D. Acemoglu, P. Restrepo, *Am. Econ. Rev.* **108**, 1488 (2018).
5. C. Goldin, L. Katz, *The Race Between Education and Technology* (Harvard Univ. Press, 2008).
6. D. H. Autor, *Science* **344**, 843 (2014).
7. E. Brynjolfsson, T. Mitchell, *Science* **358**, 1530 (2017).
8. E. Felten, M. Raj, R. Seamans, *Strateg. Manage. J.* **42**, 2195 (2021).
9. T. Eloundou, S. Manning, P. Mishkin, D. Rock, arXiv:2303.10130 [econ.GN] (2023).
10. A. Agrawal, J. S. Gans, A. Goldfarb, *The Economics of Artificial Intelligence: An Agenda* (Univ. Chicago Press, 2019).
11. K. Kanazawa, D. Kawaguchi, H. Shigeoka, Y. Watanabe, "AI, skill, and productivity: The case of taxi drivers," no. w30612, National Bureau of Economic Research (2022).
12. E. Brynjolfsson, D. Li, L. Raymond, "Generative AI at work," National Bureau of Economic Research (NBER) working paper 31161 (NBER, 2023).
13. D. H. Autor, L. Katz, M. Kearney, *Rev. Econ. Stat.* **90**, 300 (2008).
14. A. Agrawal, J. S. Gans, A. Goldfarb, *Power and Prediction* (Harvard Business Review Press, 2022).
15. E. Brynjolfsson, X. Hui, M. Liu, *Manage. Sci.* **65**, 5449 (2019).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adh9429

10.1126/science.adh9429

POLICY FORUM

Generative AI meets copyright

Ongoing lawsuits could affect everyone who uses generative AI

By **Pamela Samuelson**

Generative artificial intelligence (AI) is a disruptive technology that is widely adopted by members of the general public as well as scientists and technologists who are enthusiastic about the potential to accelerate research in a wide variety of fields. But some professional artists, writers, and programmers fiercely object to the use of their creations as training data for generative AI systems and to outputs that may compete with or displace their works (*1*, *2*). Lack of attribution and compensation for use of their original creations are other sources of aggravation to critics of generative AI. Copyright lawsuits that are now underway in the United States have substantial implications for the future of generative AI systems. If the plaintiffs prevail, the only generative AI systems that may be lawful in the United States would be those trained on public domain works or under licenses, which will affect everyone who deploys generative AI, integrates it into their products, and uses it for scientific research.

Conflicts between creators of copyrighted works and developers of technologies that enable the use of those creations in unexpected and innovative ways is nothing new. In the early 20th century, the disruptive technology of the day was player pianos. Music copyright owners sued the makers of piano rolls, claiming that rolls of their musical compositions were infringements. Subsequent copyright-disruptive technologies have included cable television, photocopiers, videotape recording machines, and MP3 players, each of which (except photocopiers) attracted copyright industry challenges (all of which failed in the courts, although Congress sometimes later extended protections in the aftermath of failed lawsuits).

Berkeley Law School, University of California Berkeley, Berkeley, CA, USA. Email: psamuelson@berkeley.edu

When new technologies pose new copyright questions that Congress did not anticipate, courts typically consider which outcome is most consistent with the constitutional purposes of copyright. The Constitution gives Congress the power "to promote the progress of Science and useful Arts," that is, to foster the creation and dissemination of knowledge for the public good. This requires balancing the legitimate interests of copyright owners to prevent misappropriations of their works that undermine incentives to create with the legitimate interests of developers of innovative technologies and follow-on creators who need some breathing space in which they, too, can innovate.

What makes generative AI more disruptive than previous technologies? One factor is certainly the exceptionally rapid pace at which generative AI technologies have been launched, adopted, and adapted. Evolution in the fields of law and policy, by contrast and of necessity, is much slower. It is, moreover, not easy to assess how to calibrate balances among competing copyright interests in the early stages of new-technology evolutions. Generative AI seems poised to have substantial impacts on the careers of professional writers and artists. During the 2023 Writers Guild of America strike, for instance, uses of generative AI are one focus of negotiations. Screenwriters are understandably worried that these technologies will displace them or diminish their compensation.

Stability AI is defending two copyright infringement lawsuits in the United States that are focused on Stable Diffusion, a widely used image generator. Getty Images is the plaintiff in one of these lawsuits. The other is a class-action lawsuit on behalf of visual artists on whose images Stable Diffusion was trained. Both complaints assert that Stability AI made unlawful copies of the plaintiffs' images when ingesting them as inputs for training Stable Diffusion's model and that output images produced by Stable Diffusion in response to user prompts are infringing derivative works.

A third generative AI lawsuit (*Doe* v. *Github, Inc.*) challenges OpenAI's development of Codex, a large language model (LLM) trained upon billions of lines of open-source software code. Also challenged is GitHub and OpenAI's collaborative development of Copilot, a coding assistant tool that draws upon the Codex LLM to suggest lines of code for specific functions in response to user prompts. (Microsoft, which owns GitHub and has invested heavily in OpenAI, is a fellow defendant.)

Rulings in favor of plaintiffs might trigger "innovation arbitrage," causing developers of generative AI systems to move their bases of operation to countries that regard the ingestion of copyrighted works as training data as fair use, like Israel's Ministry of Justice did in early 2023. Other countries that want to attract AI innovations may follow suit. If courts uphold the *Stability AI* plaintiffs' claims, OpenAI's GPT4 and Google's BARD may also be in jeopardy. Their developers would be very attractive targets of follow-on lawsuits.

INGESTING TRAINING DATA

Stability AI has yet to articulate its main defenses to the copyright charges. Insofar as the complaints allege that Stable Diffusion contains copies of in-copyright images used as training data, the claims are factually and technically inaccurate. Stable Diffusion contains an extremely large number of parameters that mathematically represent concepts embodied in the training data, but the images as such are not embodied in its model.

Training a model begins by tokenizing the contents of works ingested as training data into component elements. The model uses these tokens to discern statistical correlations—often at staggeringly large scales—among features of the content on which the model is being trained. In essence, the model is extracting and analyzing precise facts about, and correlations between, discrete elements of the works to ascertain which other discrete elements either do or do not follow or are proximate to these elements and the frequency with which the correlations do or do not exist in varying contexts.

"The complaints against Stability AI overlook the intentionally porous nature of copyrights."

The complaints against Stability AI overlook the intentionally porous nature of copyrights. What copyright law protects is only the original expression that authors contribute (such as sequences of words in a poem or the melody of music). Copyright's scope never extends to any ideas, facts, or methods embodied in works nor to elements common in works of that kind (under copyright's "scenes a faire" doctrine), elements capable of being expressed in very few ways (under the "merger" doctrine), or the underlying subjects depicted in protected works. Photographs of cats, for instance, do not give the photographer exclusive rights to characteristic features of cats, such as their noses or facial expressions. Nor does copyright's scope extend to inferences that readers might draw from reviewing an author's works, such as insights about patterns of connections among concepts or how works of that kind are constructed.

Moreover, Stability AI did not prepare the dataset on which the Stable Diffusion model was trained. This was done by a nonprofit German research organization known as LAION (Large-Scale Artificial Intelligence Open Network). LAION initially developed LAION-5B, a dataset consisting of 5.85 billion hyperlinks that pair images and text descriptions from the open internet. LAION makes this dataset available to the public for free for use as training data for those who want to use it to build generative models. LAION also developed a subset of LAION-5B, known as LAION-Aesthetics, that consists of hyperlinks to 600 million images selected by some human testers for their visual appeal and by a machine-learning analysis of human aesthetic ratings. The Stable Diffusion model was trained on the LAION-Aesthetics dataset.

LAION's creation of this dataset was very likely lawful because the European Union (EU) adopted an exemption allowing nonprofit research organizations to make copies of in-copyright works for text and data mining (TDM) purposes. The EU created this exception in recognition of the societal value of TDM as a means by which researchers can create new knowledge. This exemption cannot be overridden by contract. (A second EU exemption authorizes commercial actors to engage in TDM, although copyright owners can opt out of this exemption, as some have done.)

Stability AI makes Stable Diffusion available on an open-source basis. However, it also provides a subscription service so that those who lack resources or the inclination to host the open-source version can have access to Stable Diffusion to generate images in response to text prompts. Yet, insofar as ingesting in-copyright images to train a generative model requires making at least temporary or incidental copies of them, Stability AI is likely to argue that this is a fair use under US copyright law.

FAIR USE

Under US law, fair uses of in-copyright works do not infringe copyrights. Courts consider four factors when assessing fair use defenses: (i) the purpose of the challenged use, (ii) the nature of the copyrighted works, (iii) the amount and substantiality of the taking, and (iv) the effect of the challenged use on the market for or value of the

copyrighted work. The purpose and market effects factors are generally the most important determinants in fair-use cases, but all four factors must be weighed together in a holistic analysis.

Research, scholarship, and teaching are among the favored fair-use purposes, as are criticism, comment, and news reporting. Noncommercial uses are generally favored more than commercial uses. Since 1994, when the Supreme Court considered the fairness of 2 Live Crew's rap parody of a popular Roy Orbison song in *Campbell* v. *Acuff-Rose Music, Inc.*, courts have given considerable weight to whether the purpose of a challenged use was "transformative." The Court defined this term as uses that "add[] something new, with a further purpose or different character, altering the first with new expression, meaning, or message" [(*3*), p. 579]. Transformative uses are also less likely than nontransformative uses to harm the market for the first work. People are, for instance, unlikely to purchase 2 Live Crew's parody if they want to listen to Roy Orbison's rendition.

The *Stability AI* plaintiffs will likely argue that the ingestion of their works as training data was nontransformative and commercial. Both considerations would, if accepted, tip against fair use. However, several court decisions have ruled that analogous digital uses of in-copyright works qualified as transformative fair uses.

For example, in *Authors Guild* v. *Google, Inc.*, a court ruled that Google's digitization of millions of books from research library collections to index their contents and serve a few snippets of book contents in response to user search queries was a "highly transformative" fair use. Although Google's purpose was commercial, it was very different from the purposes for which the books were marketed. Google's use facilitated greater public access to knowledge as well as enabling TDM research and the creation of new research tools. In *Field* v. *Google, Inc.*, a court found that Google's cache copying of contents from Field's website was a transformative fair use.

The nature-of-the-work factor often has little importance in fair-use cases. The *Stability AI* plaintiffs may argue that because works of visual art lie at the very core of copyright, fair use should be thinner for these works than for the old library books at issue in the *Authors Guild* case. A countervailing consideration is that the visual artists on whose works Stable Diffusion trained made their works available on the open internet, as did Field in the *Google, Inc.*, case.

Transformative purposes tend to have spillover effects on other fair-use factors, especially the amount factor. As in the *Authors Guild* case, the *Stability AI* plaintiffs may emphasize that the defendant made exact copies of the entirety of many millions of works without permission or compensation. However, courts typically inquire whether such copying was necessary to achieve a transformative purpose. In the *Authors Guild* case, the court recognized that Google could not index book contents and serve up snippets in response to search queries unless it copied the books' contents. Stability AI will likely make a similar necessity argument about training-data usages of images.

The market effect of a challenged use is sometimes said to be the most important fair-use factor. The Getty complaint against Stablity AI emphasizes that it has established a licensing market for use of its premium photographs as training data for generative AI. That bolsters Getty's argument that Stability AI's appropriation of 12 million images from Getty websites has harmed a licensing market. The class-action claim against Stability AI is weaker because Stability AI could not have gotten a license from the class of visual artists whose works were ingested to construct the Stable Diffusion model.

The existence of a licensing market (or an intent to establish one) is not, however, a consideration that by itself can resolve a dispute in transformative fair use cases. In its 2021 *Google LLC* v. *Oracle America, Inc.*, decision, the Supreme Court rejected Oracle's argument that Google's use of parts of the Java application programming interface (API) had deprived Oracle of license revenues to which it claimed an entitlement. The Court stated that courts should consider the public benefits of a challenged use as well as potential lost revenues and how much creativity a challenged use has enabled and balance this against potential losses.

This consideration was very relevant in the *Oracle* case. Not only was Google's Android smartphone platform, in which the Java API was used, a highly innovative new software product, but it enabled millions of programmers to use their familiarity with the Java API to create many millions of programs. The Court thought this use was consistent with the constitutional objective of copyright to promote creative progress. The public greatly benefited from Android's existence and the availability of large numbers of apps that ran on that platform.

Stability AI will almost certainly channel

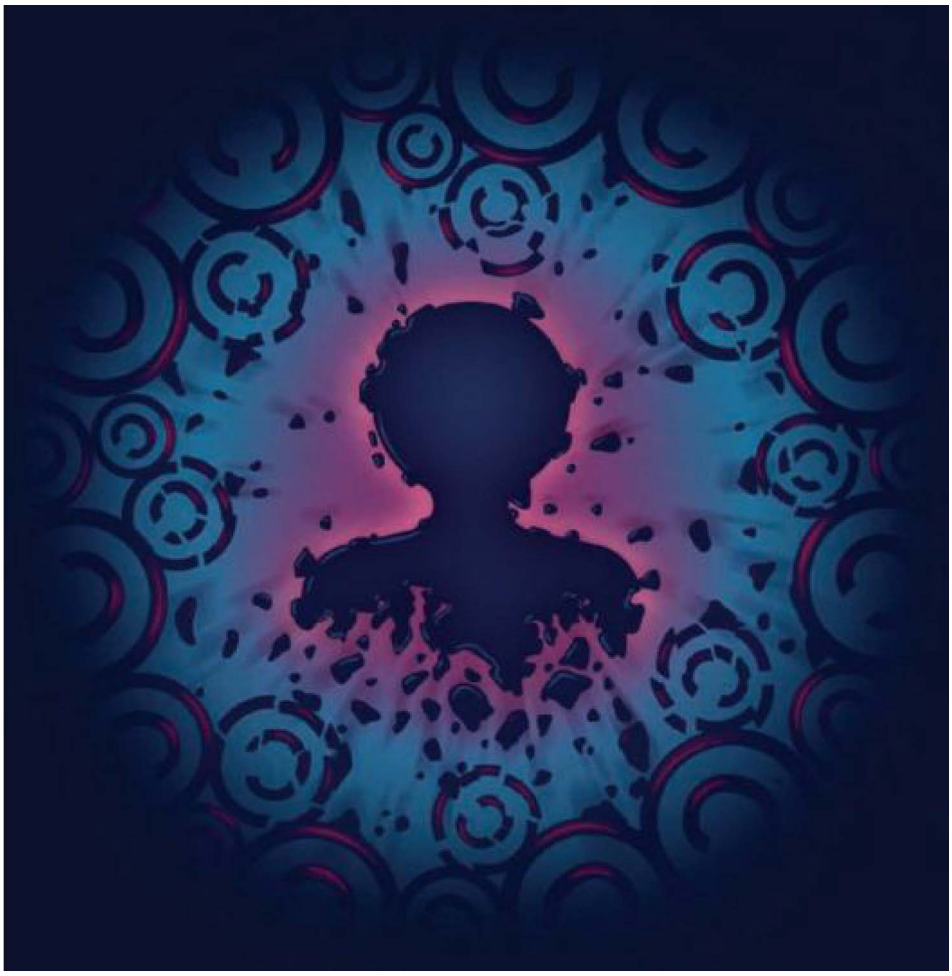

ILLUSTRATION: ANASTASIA/SCIENCE

the public-benefit and creative-impacts statements in the *Oracle* decision and point to the exceptional creativity embodied in Stable Diffusion as well as the hundreds of millions of creative uses of this generative AI system, including those by graphic artists who use it to generate ideas or refine creations.

The *Stability AI* plaintiffs will likely counter this argument with the Supreme Court's 2023 ruling in *Andy Warhol Foundation for the Visual Arts, Inc.* v. *Goldsmith*, which somewhat narrowed the conception of transformative purposes. It no longer suffices for challenged works to have a new meaning or message. More important now are whether the challenged use has a different purpose than the first work and how commercial the use is. Stability AI will argue that ingesting copyrighted materials as training data had a very different purpose than the works as first published.

What might tip the scale against Stability AI's fair-use defense is whether images produced by Stable Diffusion infringe the derivative work right of the authors of the images on which its LLM was trained. A relevant precedent is *Sega Enterprises Ltd.* v. *Accolade, Inc.*, in which an appellate court decided that Accolade had made fair use of Sega software when making reverse-engineered copies for the legitimate purpose of extracting information about how to make its videogames compatible with the Sega platform. Had Accolade reverse-engineered for an illegitimate purpose, such as to appropriate expression from the Sega games, its fair-use defense would have faltered. The Accolade games competed with Sega's games, but the court thought that this was the kind of competition among noninfringing works that copyright is supposed to foster.

OUTPUTS AS INFRINGEMENTS

The class-action complaint against Stability AI asserts that all images produced by Stable Diffusion are infringing derivative works because all are derived from the images on which its model trained. It characterizes Stable Diffusion as a "collage tool" whose outputs compete against the artists' own works and thereby harm their markets. Users of Stable Diffusion, moreover, can submit prompts requesting the generation of an image of a particular subject "in the style of" a specific named artist.

However, courts have long held that to infringe copyright's derivative work right, it is not enough to show that a second work was "based upon" an earlier work or some of its elements. The second work must have appropriated a substantial quantum of the first work's original expression. So, unless a court decides to overturn decades of precedents interpreting the derivative work right and broaden it substantially, the class action's output infringement claim is likely to fail.

The class-action complaint acknowledges that "[i]n general, none of the Stable Diffusion output images provided in response to a particular Text Prompt is likely to be a close match for any specific image in the training data" [(*4*), p. 23]. Even "in the style of" claims seem weak because copyright law does not protect styles as such. Infringement can be found only if there is a close resemblance between expressive elements of a stylistically similar work and original expression in particular works by that artist.

The reason that Stable Diffusion outputs are highly unlikely to be substantially similar to particular images on which its model was trained is due to how Stable Diffusion assembles them. Constructing a model for an image-generating AI requires processing enormous quantities of input data to produce abstract representations of image elements (such as cats playing with a ball on a linoleum floor). Diffusion adds noise to image elements when encoding them. The pairing of text descriptions and images allows the model to cluster the abstract representations so that similar representations will be in proximity (representations of cats near other cat representations). When a user enters a prompt directing the software to generate a specific type of output, the generative AI system uses complex statistical calculations to assemble an output that the system predicts will match what the user requested.

It is, however, possible for generative AI outputs to infringe copyrights. If the same input image (say, of Mickey Mouse) is present in many works on which the model was trained and its developer did not follow industry best practices by eliminating duplicates and using output filters to prevent infringements, user prompts could result in infringing outputs (although this user, not the developer of the generative AI system, may be the infringer). Ironically, the larger and more diverse that the dataset on which a generative model was trained is, the less likely are infringing outputs.

The Getty complaint against Stability AI is more modest in its infringing output claims. Yet Getty, too, may find it difficult to prove that particular Stable Diffusion outputs are substantially similar to particular photographs to which it owns copyrights. In general, Stable Diffusion outputs will be distinguishably different from the images on which the model was trained.

The *Stability AI* plaintiffs will likely emphasize that the images produced by Stable Diffusion compete with their works in the marketplace. They can point to the Supreme Court's *Goldsmith* decision, which treated competing uses as weighing against fairness. Yet *Goldsmith* involved two works that were substantially similar in their expressions—Goldsmith's photo of Prince and Warhol's print derived from Goldsmith's photo—that competed in the same licensing market for magazines. Stability AI will be relying on differences in Stable Diffusion's outputs relative to plaintiffs' works in order to distinguish their case's context from that of *Goldsmith*.

CONCLUDING THOUGHTS

Based on existing precedents and an understanding about how Stable Diffusion was trained and how it generates images in response to prompts, Stability AI seemingly has a reasonable chance of prevailing on the copyright claims. (Both the Getty and the class-action complaints raise other claims that cannot be addressed in this brief article.) The lawsuits are, however, in very early stages, and it may be years before courts render decisions.

In mid-May 2023, Congress held its first hearing about generative AI and copyright issues, during which witnesses expressed divergent views. The US Copyright Office is well aware of the consternation that generative AI has fomented in copyright-dependent communities. The Office hosted "listening sessions" in spring 2023 to provide stakeholders with opportunities to explain their perspectives on the two principal questions posed in the *Stability AI* cases: Is the use of in-copyright works as training data for generative AI systems an infringement of copyright? Are the outputs of generative AI systems infringing derivative works?

During the summer of 2023, the Office plans to allow interested parties to submit written comments expressing their perspectives and analyses on these and related questions. The Office intends to write a report setting forth its conclusions, which may include legislative recommendations. Scientists who have an interest in the future of generative AI would be well advised to submit comments. ■

REFERENCES AND NOTES

1. N. Klein, "AI machines aren't 'hallucinating.' But their makers are," *The Guardian*, 8 May 2023.
2. M. Cuenco, "We must declare jihad against AI," *Compact*, 28 April 2023.
3. *Campbell* v. *Acuff-Rose Music, Inc.*, 510 US 569, 579 (1994).
4. Complaint, *Anderson* v. *Stability AI, Ltd.*, case no. 3:23-cv-00201 para. 93 (N.D. Cal. Jan. 13, 2023).

10.1126/science.adi0656